UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/723,049	11/26/2003	Xian-Sheng Hua	MS1-1744US	7823

22801   7590   11/21/2007
LEE & HAYES PLLC
421 W RIVERSIDE AVENUE SUITE 500
SPOKANE, WA 99201

EXAMINER
GISHNOCK, NIKOLAI A

ART UNIT	PAPER NUMBER
3714	

MAIL DATE	DELIVERY MODE
11/21/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

PTOL-90A (Rev. 04/07)

	Application No.	**Applicant(s)**
Office Action Summary	10/723,049	HUA ET AL.
	Examiner	**Art Unit**	
	Nikolai A. Gishnock	3714	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on <u>06 September 2007</u>.

2a)☒ This action is **FINAL**.   2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-15 and 17-44* is/are pending in the application.

     4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-15 and 17-44* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on <u>26 November 2003</u> is/are: a)☒ accepted or b)☐ objected to by the Examiner.

     Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

     Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

     a)☐ All   b)☐ Some *   c)☐ None of:

         1.☐ Certified copies of the priority documents have been received.

         2.☐ Certified copies of the priority documents have been received in Application No. _____.

         3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

     * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
     Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413)
     Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

In response to Applicant's remarks filed 9/6/2007, claim 16 is cancelled. Claims 1-15 & 17-44

are pending.

Claim Rejections - 35 USC § 103

1.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

2.   The factual inquiries set forth in *Graham* v. *John Deere Co.*, 383 U.S. 1, 148 USPQ 459

(1966), that are applied for establishing a background for determining obviousness under 35

U.S.C. 103(a) are summarized as follows:

1.   Determining the scope and contents of the prior art.
2.   Ascertaining the differences between the prior art and the claims at issue.
3.   Resolving the level of ordinary skill in the pertinent art.
4.   Considering objective evidence present in the application indicating obviousness
        or nonobviousness.

3.   Claims 1, 2, 8, 20, 23-25, 27-29, 40, & 41 are rejected under 35 U.S.C. 103(a) as being

unpatentable over Stelovsky (US 5,782,692), hereinafter known as Stelovsky, in view of Wang,

(US 2002/0133764 A1), hereinafter known as Wang.

4.   Stelovsky teaches a processor-readable medium comprising executable instructions for

personalizing karaoke (Column 1, Lines 54-67), comprising: segmenting visual content to

produce a plurality of sub-shots and segmenting music to produce a plurality of music sub-clips

(multimedia presentation track consisting of video, audio, and text display is segmented with ·

respect to specific beginning and ending points, Column 3, Lines 27-65); and displaying at least some of the plurality of sub-shots as a background to lyrics associated with the plurality of music sub-clips ("Karaoke Game" presentation has synchronized video and instrumental sound tracks, Column 9, Lines 15-21; the text can be superimposed on the video, Column 10, Lines 5-6) [Claims 1 & 8].

5.   Stelovsky teaches a processor-readable medium comprising instructions for providing lyrics for integrating lyrics, music, and video content suitable for karaoke, comprising instructions for: receiving a request for a file associated with a specific song (clicking on a word in the text track, Column 14, Lines 42-48), wherein the file comprises music, lyrics, and timing values (The time-dependent sequence is composed of tracks that are synchronized with respect to a common time axis {hereinafter "multimedia presentation"}. The basic track consists of video display images and is synchronized with at least one other track that consists of audio or text display, 3:31-35; The multimedia presentation is segmented with respect to specific beginning and ending points of segments on the time axis, i.e. there are one or more points of time that partition the time axis into time segments, 3:52-55), and fulfilling the request by sending the file associated with the specified song (connection is established with a remote on-line service, search query initiated, and results are displayed, Column 14, Lines 42-48), segmenting visual content to produce a plurality of sub-shots of a length corresponding to the music sub-clips (multimedia presentation track consisting of video, audio, and text display is segmented with respect to specific beginning and ending points, Column 3, Lines 27-65), and outputting the plurality of music sub-clips together with corresponding sub-shots of visual content, which is configured as a background to the lyrics associated with the music sub-clips ("Karaoke Game" presentation has synchronized video and instrumental sound tracks, Column 9, Lines 15-21; the text can be superimposed on the video, Column 10, Lines 5-6) [Claim 23].

6.   Stelovsky teaches a personalized karaoke device, comprising: a music analyzer

configured to create music sub-clips of varying lengths according to a song (Segmentation

Authoring System {SAS} facilitates the identification of points in time where a segment starts

and ends, Column 5, Line 62 to Column 6, Line 2; multimedia presentation track consisting of

video, audio, and text display is segmented with respect to specific beginning and ending points,

Column 3, Lines 27-65); a visual content analyzer configured to define and select visual content

sub-shots (Using SAS, the author partitions the multimedia presentation into time segments

according to predominant time units, e.g., measures of song, sound bites, or action sequences

in a movie, Column 6, Lines 51-54); a lyric formatter configured to time delivery of syllables of

lyrics of the song (evaluation feedback of user's input includes visualization of differences in

pronunciation patterns, processes involved in generating {human} speech, such as positions of

the tongue and air flow patterns, Column 14, Lines 52-59; it is inherent that the speech analysis

as disclosed could recognize syllables and sentences, which are pronunciation patterns);

sections of the text track are linked to the time segments, Column 6, Line 55); and a composer

configured to assemble the music sub-clips with the visual content sub-shots, and configured to

adjust the length of the sub-shots to correspond to the music sub-clips, and to superimpose the

syllables of the lyrics of the song over the sub-shots ({SAS} sections of a text track and

additional media resources are linked to the time segments, Column 6, Lines 55-57) [Claim 25].

7.   Stelovsky teaches an apparatus, comprising: means for creating music sub-clips

according to a song, and means for defining and selecting visual content sub-shots (multimedia

presentation track consisting of video, audio, and text display is segmented with respect to

specific beginning and ending points, Column 3, Lines 27-65); means for timing delivery of

syllables of lyrics of the song (sections of the text track are linked to the time segments, Column

6, Line 55; the text can be superimposed on the video, Column 10, Lines 5-6, also Column 14,

Lines 52-59 and Column 9, Lines 15-21); and means for assembling the music sub-clips with

the visual content sub-shots, adjusting the length of the sub-shots to correspond to the length of

the music sub-clips, and to superimpose the syllables of the lyrics of the song over the sub-

shots [Claim 40].

8.   What Stelovsky fails to teach is where the segmenting of music to produce a plurality of

music sub-clips establishes boundaries between the music sub-clips at beat positions within the

music [Claims 1, 23, 25, & 40], wherein a music analyzer is configured to segment the song with

a beat between each of the music sub-clips [Claim 27], and wherein each sub-clip has a

duration that is a function of song tempo [Claim 28]. However, Wang teaches a method of

detecting beats in a music stream (Beat is defined in the relevant art as a series of perceived

pulses dividing a musical signal into intervals of approximately the same duration. Beat

detection can be accomplished by any of three methods. The preferred method uses the

variance of the music signal, which variance is derived from decoded Inverse Modified Discrete

Cosine Transformation (IMDCT) coefficients. The variance method detects primarily strong

beats. The second method uses an Envelope scheme to detect both strong beats and offbeats.

The third method uses a window-switching pattern to identify the beats present. The window-

switching method detects both strong and weaker beats. In one embodiment, a beat pattern is

detected by the variance and the window switching methods. The two results are compared to

more conclusively identify the strong beats and the offbeats, Para. 0070-0074; see also Figure

7, the numbered delta functions are understood to be detected beats), and segmenting the

music stream at beat boundaries (A normal, error-free audio transmission is represented in the

top graph {of Figure 6} by a first and second beat-to-beat interval waveform. The first waveform

includes a first beat and the audio information up to a second beat. Similarly, the second

waveform includes the second beat and the audio information up to a third beat; In accordance

with the method of the present invention, a replacement waveform, including a replacement

beat, is copied from the first beat and the first waveform; and is substituted for the missing audio

segment in the time interval T_1 to T_2, as shown in the bottom graph; all at Para. 0058-0069; see

also Figure 6). The beat intervals are taught by Wang to be a function of song tempo (the beat-

to-beat interval is replaced by the audio data frames from a corresponding beat-to-beat interval

in a preceding 4/4 bar. Most popular music has a rhythm period in 4/4 time, Para. 0067; 4/4 time

is understood to be a tempo). Any of the three methods taught by Wang would be used to

detect beats in a music clip, and Wang's method of copying and pasting music waveforms

segmented by at beat positions would be used to align video, still pictures, music, and lyrics

along those boundaries, in the manner as taught by Stelovsky. Therefore, it would have been

obvious to one of ordinary skill in the art, at the time the invention was made, to have used

Wang's methods of segmenting of music to produce a plurality of music sub-clips, establishing

boundaries between the music sub-clips at beat positions within the music, with the methods of

Stelovsky for integrating lyrics, music, and video content suitable for karaoke, in order to exploit

the beat pattern of music signals to improve the presentation of music when transferred over a

network [Claims 1, 23, 25, 27, 28, & 40].

9.   Stelovsky teaches instructions for shortening some of the plurality of sub-shots to a

length of a corresponding music sub-clip (the system displays the current segment's start and

end points, so the author can select and edit the boundary points, Column 7, Lines 14-19)

[Claim 2].

10.   Stelovsky teaches instructions for: obtaining lyrics from a file (textual track can be

generated remotely and transmitted using communications means, Column 14, Lines 20-24);

and coordinating delivery of the lyrics with the music using timing information contained within

the file (Column 3, Lines 52-65) [Claim 20].

11.   Stelovsky teaches wherein obtaining lyrics comprises instructions for sending the file

over a network to a karaoke device (textual track can be generated remotely and transmitted

using communications means, Column 14, Lines 20-24; on-line services provide downloading of

files, e.g. Internet, Column 6, Lines 49-50) [Claim 24].

12.   Stelovsky teaches wherein the visual content analyzer is configured to segment video

into sub-shots (Column 6, Lines 51-54) [Claim 29].

13.   Stelovsky teaches wherein the means for defining and selecting visual content sub-shots

is a video analyzer configured to segment video into sub-shots (Using SAS, the author partitions

the multimedia presentation into time segments according to predominant time units, e.g., action

sequences in a movie, Column 6, Lines 51-54) [Claim 41].

14.   Claim 3 is rejected under 35 U.S.C. 103(a) as being unpatentable over Stelovsky, in

view of Golin (US 5,990,980), hereinafter known as Golin. Stelovsky teaches all the features as

described above in the rejection of claim 1. What Stelovsky fails to teach is wherein segmenting

the visual content comprises instructions for: dividing a shot into two sub-shots at a maximum

peak of a frame difference curve; and repeating the dividing to result in sub-shots shorter than a

maximum sub-shot length. However, Golin teaches the use of a Frame Dissimilarity Measure

(FDM), which is the ratio of a net dissimilarity measure and a cumulative dissimilarity measure

of two consecutive frames (Column 3, Line 65 to Column 4, Line 12). The processing of sub-

shots uses the FDM to identify transitions between shots in a video sequence, which appear as

peaks in the FDM data (Column 5, Lines 21-42). The data analysis for the sub-shot dividing is a

loop, which starts with frames at the beginning of the video sequence and scans through the

data to the frames at the end of the sequence (Column 5, Lines 54-62). The length of the entire

video sequence is a maximum sub-shot length. Therefore, it would have been obvious to one of

ordinary skill in the art, at the time the invention was made, to have used the FDM peak analysis of dividing sub-shots, as described in Golin, for the video segmenting used in Stelovsky, in order to more effectively detect gradual transitions between subshots [Claim 3].

15.   Claims 4-7, 32, & 33 are rejected under 35 U.S.C. 103(a) as being unpatentable over Stelovsky, in view of Osberger (US 6,670,963), hereinafter known as Osberger. Stelovsky teaches all the features as described above in the rejection of claims 1 & 25. What Stelovsky fails to teach is wherein the filtering of a plurality of sub-shots is according to importance or quality [Claim 4]. However, Osberger teaches giving areas of medium motion high importance (Column 7, Lines 10-21). Osberger also teaches that areas of low texture (quality) such as faces are strong attractors of attention (Column 8, Lines 40-54). The sub-shots that are high in "regions of interest", or attention attracting, are identified (filtered) as taught by Osberger (Column 2, Lines 24-41). Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made, to have used the methods of Osburger for filtering sub-shots based on attention indices such as importance to the camera and texture quality, in the karaoke video segmenting device of Stelovsky, in order to increase the entertainment value of the karaoke experience to a user [Claim 4]. What Stelovsky also fails to teach is wherein filtering the plurality of sub-shots according to importance comprises instructions for evaluating frames within a sub-shot according to attention indices, and averaging the attention indices for the frames to determine if the sub-shot should be included [Claim 6]. However, Osberger teaches identifying and adaptively segmenting frames of video based upon an attention model, AKA total importance map, composed by linear weighting of the spatial and temporal importance maps (Column 2, Lines 24-41). It is inherent that averaging is merely linear weighting with a weight factor of one. Therefore, it would have been obvious to one of ordinary skill in the art, at the time

the invention was made, to have utilized the averaging of the attention indices of Osberger to select frames of importance, for use in the karaoke system of Stelovsky, in order to adapt the attention model for a variety of different types of video sub-shots, while accurately determining regions of interest in the videos [Claim 6]. What Stelovsky also fails to teach is wherein filtering the sub-shots according to importance comprises instructions for analyzing the camera motion, object motion, and specific objects within the subshots, and filtering the subshots according to the analysis [Claim 7], or wherein a visual content analyzer is configured to select from the sub-shots according to ranked importance, gauged by detection of color entropy, object motion, camera motion, or of a face within the sub-shot [Claim 32]. However, Osberger teaches selecting or filtering sub-shots by color information (Column 3, Lines 6-15), by camera or object motion (Column 7, Lines 7-37), or by specific objects, including faces, in a sub-shot (Column 8, Lines 40-54). Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made, to have used the various color, motion, and object detection in the video sub-shots, as described by Osberger, in the personalized karaoke system on Stelovsky, in order to improve the prediction of visual importance of a sub-shot [Claims 7 & 32]. What Stelovsky further fails to teach is wherein filtering the plurality of sub-shots comprises instructions for: examining color entropy within each of the plurality of sub-shots to detect motion more than a threshold indicating interest and less than a threshold indicating low camera and/or object movement; and selecting sub-shots having acceptable motion and/or color entropy scores [Claim 5], or wherein the visual content analyzer is configured to filter out sub-shots having low image quality, as measured by low entropy and low motion intensity [Claim 33]. However, Osberger teaches segmenting frames into regions based upon both color and luminance (Column 2, Lines 24-41). The term entropy is taken to mean Information Entropy or Shannon Entropy, which refers to a measure of uncertainty associated with a random variable.

Thus, referring to lossless data compression, the color entropy would refer to an average

minimum number of bits needed to communicate a color value. Osberger teaches using an

algorithm to segment an image into homogeneous regions using color information, to generate

the spatial importance map (Column 3, Lines 6-15). Osberger also teaches that, if the spatial

importance map is too noisy from frame to frame, a temporal smoothing operation is performed,

and a temporal importance map is generated (Column 6, Line 66 to Column 7, Line 37). The

temporal importance map is calculated using adaptable thresholds because the amount of

motion varies greatly across different scenes. Osberger also teaches identifying sub-shots with

regions of interest by using the spatial and temporal interest maps in order to produce an

adaptive segmentation model (Column 8, Lines 58-67), for segmenting video scenes. Therefore,

it would have been obvious to one of ordinary skill in the art, at the time the invention was made,

to have incorporated the color entropy detection, then the camera motion detection of Osberger

with the segmentation of karaoke video as described by Stelovsky, in order to attract the interest

of a karaoke user more effectively [Claims 5 & 33].

16.   Claims 9-11 & 34 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Stelovsky, in view of Osberger, as applied to claims 1, 8, & 25 above, and further in view of

Paniconi et al. (US 2007/0064806 A1), hereinafter known as Paniconi. Stelovski and Osberger

teach all the features as shown above in the rejections of claims 1, 8, and 25 above. Osberger

teaches selecting important sub-shots from within the plurality of sub-shots [Claim 9], evaluating

color entropy, camera motion, and object motion, and detecting objects, and selecting the

important sub-shots based on the·evaluation [Claim 10]; and a visual content analyzer

configured to select sub-shots of a greater importance [Claim 34]. What Stelovsky and Osberger

fail to explicitly teach is wherein the sub-shots are uniformly distributed over the run-time of a

source video [Claims 9, 11, & 34], or evaluating normalized entropy of the sub-shots along a

time line of video from which the sub-shots were obtained [Claim 11]. However, Paniconi

teaches filtering video images by distinguishing a uniform pattern of motion vectors, evenly

distributed across the target images in a video compression scheme [Para. 0016-0018]. It is

inherent that the filtering prediction is an attention model because, in any lossy compression

scheme, frames of high importance are retained in order to convey the video information while

the least important frames are discarded. Paniconi also teaches normalizing the motion vectors

(low pass filter, Para. 0037-0040). Normalization of data can be described as the process of

removing statistical errors in data. A low pass filter removes motion error, thus normalizing the

entropy of video data. Therefore, it would have been obvious to one of ordinary skill in the art, at

the time the invention was made, to have selected important yet uniformly distributed sub-shots

by evaluating normalized entropy as in Paniconi, in light of the importance indices of Osberger,

in the karaoke system of Stelovsky, for the purpose of maximizing the average importance of

the video sub-shot while minimizing the extraneous frames of less importance by filtering

[Claims 9-11 & 34].

17.   Claims 12-15, 31, 36, 37, & 43 are rejected under 35 U.S.C. 103(a) as being

unpatentable over Stelovsky, in view of Geigel et al. (US 2002/0122067 A1), hereinafter known

as Geigel. Stelovsky teaches all the features as demonstrated in the rejection of claims 1, 25, &

40 above. What Stelovsky fails to explicitly teach is wherein the instructions for segmenting

visual content includes assigning photographs to be sub-shots [Claim 12], wherein the visual

content comprises home video and photographs in digital formats [Claim 15], or wherein a

visual content analyzer is configured to assemble still photographs, each of which is a sub-shot

[Claim 31], and instructions for assigning photographs includes converting at least one

photograph to video [Claim 14]. However, Geigel teaches a layout generator for digital images (Para. 0010), including photographs or video clips (Para. 0055), which converts the images into a video (output is Picture CD media or other photo delivery media, Para. 0057). It is inherent that a series of images displayed during a progression of time is a video. Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made, to have assembled and converted photos to video, as taught by Geiger, for the background video in the entertainment system of Stelovsky, in order to automate the layout of the background in a manner pleasing to the user [Claims 12, 14, 15, & 31]. What Stelovsky also fails to teach is wherein a visual content analyzer is configured with instructions for assigning photographs includes instructions for: rejecting photographs having problems with quality [Claim 13]; and rejecting a similar group of photographs when one within the group has been selected [Claims 13 & 37]. However, Geigel teaches performing detection of dud images and duplicate images prior to being submitted to the layout system (Para. 0061). Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made, to have not selected dud or duplicate images when creating the background image layout, as shown by Geigel, when implementing the entertainment system of Stelovsky, in order to necessitate the minimal input from the user when assembling images aesthetically pleasing to the user [Claims 13 & 37]. What Stelovsky further fails to teach is wherein a visual content analyzer is configured to organize photographs by the date of exposure and scene, thereby obtaining photographs having a relationship [Claim 36]. However, Geigel teaches organizing the images (page layout algorithm, Para 0059) by date of exposure (chronology of the images, Para. 0063) and scene (event clustering, Para. 0060). It is inherent that all the photographs would thus be related by a date range or event group. Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made, to have organized the images to the extent provided by

Geigel, is the operation of the entertainment system of Stelovsky, in order to distribute the

photographs automatically according to an algorithm that valued a user-pleasing arrangement

[Claim 36].

18.   Claims 17 & 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Stelovsky. Stelovsky teaches all the features as demonstrated in the rejection of claim 1 above.

What Stelovsky fails to explicitly teach is wherein the segmenting music comprises instructions

for bounding the sub-clip's length according to: minimum length = min(max(2*tempo,2),4) and

maximum length = minimum length+2 [Claim 17], or establishing the music sub-clip's length

within a range of 3 to 5 seconds [Claim 18]. However, Applicant has not disclosed that having

(min(max(2*tempo,2),4) < length < min(max(2*tempo,2),4)+2) or (3 < length < 5) seconds

solves any stated problem or is for any particular purpose. Moreover, it appears that the

arbitrary length of the sub-clips of Stelovsky or the Applicant's instant invention would perform

equally well for synchronizing the sub-clips with a video. Accordingly, it would have been

obvious to one of ordinary skill in the art, at the time the invention was made, to have modified

Stelovsky such that the music sub-clips had a rigid minimum and maximum length, because

such a modification would have been considered a mere design consideration, which fails to

patentably distinguish over Stelovsky [Claims 17 & 18].

19.   Claims 19, 39, & 44 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Stelovsky, in view of Bloom et al. (US 2005/0042591 A1), hereinafter known as Bloom.

Stelovsky teaches all the features as demonstrated above in the rejections of claims 1, 18, 25, &

40 above, including wherein the lyric formatter is configured to consume a file detailing timing of

the lyrics (the textual track can be generated remotely and transmitted by communication

means, digitally, using a software program, Column 14, Lines 14-24; the digital textual track used for the karaoke is inherently a file to be "consumed" or used). Stelovsky teaches wherein evaluation of output can involve differences in pronunciation patterns and any processes involved in generating speech (Column 14, Lines 52-59). What Stelovsky fails to teach is wherein segmenting the music comprises a lyric formatter configured with instructions for establishing boundaries for the music sub-clips at sentence breaks [Claim 19], or consuming a file detailing timing of each syllable and each sentence of the lyrics [Claims 39 & 44], and for rendering the lyrics syllable by syllable [Claim 44]. However, Bloom teaches automatically synchronizing sound to images, wherein lyric segmentation may be syllable by syllable (line can be a single word or sound) or a sentence (Para. 0139). Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made, to have segmented the music of the karaoke system of Stelovsky, in light of the syllable and sentence boundaries of the lyrics as taught by Bloom, in order to synchronize the song with a user's lip movements on the accompanying video display [Claims 19, 39, & 44].

20.   Claim 21 is rejected under 35 U.S.C. 103(a) as being unpatentable over Stelovsky, in view of Tsai (US 6,572,381 B1), hereinafter known as Tsai. Stelovsky teaches all the features as demonstrated above in the rejections of claims 1 & 20 above. What Stelovsky fails to teach is wherein obtaining the lyrics comprises instructions for sending the file over a network to a karaoke device as part of a pay-for-play service [Claim 21]. However, Tsai teaches a plurality of karaoke terminals connected to a host computer via a network (communications line) that delivers lyric data (Column 8, Lines 48-61). Tsai teaches a karaoke system shares the source data as part of a pay service (Column 2, Lines 48-56; also Column 20, Line 52 to Column 21, Line 56). Therefore, it would have been obvious to one of ordinary skill in the art, at the time the

invention was made, to have sent the lyrics file over a network in conjunction with a pay-for-play service, as taught by Tsai, in the karaoke system of Stelovsky, in order to offer commercial messages with updated custom content to a subscriber of a karaoke service [Claim 21].

21.   Claim 22 is rejected under 35 U.S.C. 103(a) as being unpatentable over Stelovsky, in view of Tashiro et al. (US 5,703,308), hereinafter known as Tashiro. Stelovsky teaches all the features as demonstrated above in the rejections of claim 1 above. What Stelovsky fails to teach is wherein the processor-readable medium comprises instructions for: querying a database of songs by humming a portion of a desired song; and selecting the desired song from among a number of possibilities suggested by an interface to the database [Claim 22]. However, Tashiro teaches a karaoke device having database of songs (music data storage device with a plurality of entry songs stored in a data table, Column 1, Line 54 to Column 2, Line 3), wherein the database is queried by humming a song (key melody patterns which represent a desired song are input by voice, Column 3, Lines 10-14) and selecting the desired song through an interface (music selection is made from top 10 matching entries, Column 7, Lines 48-67). Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made, in the karaoke system of Stelovsky, to search and select a desired song from a database by humming, as taught by Tashiro, in order to select a song even if neither the artist nor the title of the song is known [Claim 22].

22.   Claim 26 is rejected under 35 U.S.C. 103(a) as being unpatentable over Stelovsky, in view of Trovato et al. (US 7,058,889 B2), hereinafter known as Trovato. Stelovsky teaches all the features as demonstrated above in the rejections of claims 1 & 25. What Stelovsky fails to teach wherein the music analyzer is configured to segment the song with a strong onset

between each of the music sub-clips [Claim 26]. However, Trovato teaches locating transition

points for a music segmentation scheme by onset break detection (Column 7, Lines 33-51; also

Figure 6). It is inherent from Figure 6 that weak onset breaks are not used as transition points.

Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention

was made, to have analyzed the music used in the karaoke system of Stelovsky with the onset

break detection method defined in Trovato, in order to automatically synchronize the music with

the background video consistent with human perception [Claim 26].

23.   Claim 27 is rejected under 35 U.S.C. 103(a) as being unpatentable over Stelovsky, in

view of Kondo (US 6,232,540 B1), hereinafter known as Kondo. Stelovsky teaches all the

features as demonstrated above in the rejections of claims 1 & 25. What Stelovsky fails to teach

is wherein a music analyzer is configured to segment the music automatically, comprising

instructions for: establishing boundaries for the music sub-clips with a beat position between

each of the music sub-clips [Claim 27]. However, Kondo teaches establishing boundaries

(positions) for music sub-clips (rhythm sound source signals) at beat positions within the music

(positions of attacks in the rhythm sounds, Abstract). Therefore, it would have been obvious to

one of ordinary skill in the art, at the time the invention was made, to have divided the music

sub-clips at beat positions within the music, as shown in Kondo, for use in the karaoke system

of Stelovsky, in order to avoid occurrences of rhythm disorder in the rhythm sounds [Claim 27].

24.   Claims 30 & 42 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Stelovsky, in view of Borden, IV et al. (US 2003/0200105 A1), hereinafter known as Borden IV.

Stelovsky teaches all the features of claims 25 & 40 above. What Stelovsky fails to teach is a

video analyzer or visual content analyzer configured to access folders of home video and

photographs containing content from which the sub-shots are derived [Claims 30 & 42].

However, Border IV teaches a video analyzer (user's data processing device) which can access

folders of a customer's video or photographs (MY PHOTOS homepage document, containing a

user's uploaded images or video, Para. 0016-0017). Therefore, it would have been obvious to

one of ordinary skill in the art, at the time the invention was made, to have accessed a user's

personal video and photo content for generating the sub-shots, in the karaoke device of

Stelovsky, in order to attract potential customers to receive services by hosting their personal

data [Claims 30 & 42].

25.   Claims 35, 38, & 43 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Stelovsky, in view of Osberger, as applied to claims 25 & 40 above, and further in view of

Geigel. Stelovsky teaches all the features of claims 25 and 40 above. What Stelovsky fails to

teach is wherein a visual content analyzer is configured to reject photographs of low quality by

detecting over and under exposure, overly homogeneous images, and blurred images [Claim

35]. Osberger teaches a visual analyzer (image processing algorithm) to detect overexposure

and underexposure (contrast), overly homogeneous images (homogeneous regions, Column 3,

Lines 6-15), and blurred images (areas of very high motion, Column 7, Lines 10-26). What

Stelovsky and Osberger fail to teach is wherein the visual content analyzer rejects photographs

which are underexposed, overexposed, overly homogeneous, or blurred [Claim 35]. However,

Geigel teaches selection of the best image (Para. 0057). Therefore, it would have been obvious

to one of ordinary skill in the art, at the time the invention was made, to have rejected images

which are underexposed, overexposed, overly homogeneous, or blurred, in light of the

teachings of Osberger and Geigel, in the entertainment system of Stelovsky, in order to

discriminate images to present highly desirable visuals to a karaoke user [Claim 35]. What

Stelovsky further fails to teach is wherein the means for defining and selecting visual content

sub-shots is a video analyzer configured for: detecting an attention area within a photograph;

and creating a photo to video sub-shot based on the attention area, wherein the video includes

panning and zooming [Claims 38 & 43]. Osberger teaches a visual analyzer (image processing

algorithm) to detect an attention area within a photograph (Column 2, Lines 24-41), and wherein

motion vectors are used by camera motion estimation algorithm to determine pan and zoom in a

frame (Column 7, Lines 22-37). What Stelovsky and Osberger fail to teach is wherein photo to

video subshot includes panning and zooming. However, Geigel teaches, in photography terms

rather than videography terms, panning the images (auto-cropping, Para. 0057) and zooming

the images (scaling, Para. 0122). Therefore, it would have been obvious to one of ordinary skill

in the art, at the time the invention was made, to created a photo to video sub-shot based on a

detected attention area, including panning and zooming, in light of the teachings of Osberger

and Geigel, in the entertainment system of Stelovsky, in order to further refine the content

information of an image by focusing on the attention-attracting elements in the photo to video,

when used as the background for karaoke entertainment [Claims 38 & 43].

Response to Arguments

26.   Applicant's arguments, filed 9/6/2007, see pages 14-19, with respect to claims 1, 23, 25,

& 40, have been considered but are moot in view of the new ground(s) of rejection. Applicant's

arguments are further not responsive because the rejection of claims 16 & 27, regarding the

limitation of detecting beat positions within the music, and establishing boundaries at the beat

positions, in view of the Kondo (US 6,232,540 B1) reference was not addressed. See Office

Action of 4/6/2007, pages 11-12, paragraph 16.

Conclusion

27.   Applicant's amendment necessitated the new ground(s) of rejection presented in this

Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant

is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within TWO

MONTHS of the mailing date of this final action and the advisory action is not mailed until after

the end of the THREE-MONTH shortened statutory period, then the shortened statutory period

will expire on the date the advisory action is mailed, and any extension fee pursuant to 37

CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event,

however, will the statutory period for reply expire later than SIX MONTHS from the date of this

final action.

Any inquiry concerning this communication or earlier communications from the examiner

should be directed to Nikolai A. Gishnock whose telephone number is 571-272-1420. The

examiner can normally be reached on M-F 8:30a-5p.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Xuan M. Thai can be reached on 571-272-7147. The fax phone number for the

organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent

Application Information Retrieval (PAIR) system. Status information for published applications

may be obtained from either Private PAIR or Public PAIR. Status information for unpublished

applications is available through Private PAIR only. For more information about the PAIR

system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private

PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you

would like assistance from a USPTO Customer Service Representative or access to the

automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

NAG

NAG
11/16/2007

Ronald Laneau
Primary Examiner
Art Unit 3714

11/17/07